Chapter 5

Case study

5.1 Introduction

It was stated earlier in this thesis that a useful contribution to the field of reliability modeling can only be claimed if the developed theory is implemented successfully on an actual data set. Data can be obtained in two ways: (i) it can be generated in a laboratory under controlled conditions; or (ii) it can be collected in a typical industrial situation. Successful application of the theory on data obtained from a laboratory may be doubtful because of the controlled conditions in laboratories that do no exist in practice. It was hence decided to use data from a real industrial situation.

Data satisfying the requirements of proportional intensity modeling was found at SASOL Coal's Twistdraai plant at Secunda*. The Twistdraai plant was started up in September 1996 as a coal washing plant that washes coal to a certain cleanliness before it enters the petrochemical process. In the plant, eight Warman® axial in, radial out pumps are used to circulate a water and magnetite solution which is used in the washing process. A condition monitoring maintenance strategy through vibration monitoring was applied on the pumps from the startup date. Despite this strategy, to date several failures have occurred on these pumps. The events produced by the pumps together with the vibration information are used in the PIM theory of Chapter 3.

The theory that was developed in this thesis not only needs to be applied to real data but also to be benchmarked against existing approaches. The only existing approach that the present research can be compared against is the decision technique of Makis and Jardine mentioned in Section 1.4.2. This approach was applied to the above-mentioned data set by Vlok (1999) and the results are briefly repeated in this chapter and compared to the RLE approach.

*SASOL is a major petrochemical company in South Africa

Chapter 5 starts off with a description of the SASOL data set after which the approach of Makis and Jardine is applied and discussed. Because this approach was only briefly introduced in Chapter 1, more details on the theory are also presented for completeness. The second part of this chapter consists of the application of the RLE theory developed in this thesis on the same data set. Chapter 5 ends with a detailed comparison of the two approaches.

5.2 Description of SASOL data

The data set under discussion has many shortcomings, including missing observations and irregular inspection intervals, but was the best data set found after an extensive search for suitable data in the South African industry. The Twistdraai plant was started up in September 1996 and is thus still relatively new. Data was collected from September 1, 1996 to November 1, 1998 which gives an analysis time horizon of 791 days. A second data set was collected from November 1, 1998 to February 28, 1999 to further evaluate the combined PIM's performance.

5.2.1 Background

A total of eight identical axial in, radial out, Warman pumps are used in a specific section of the plant to circulate a water and magnetite solution. These pumps are very important in the washing process and significant production losses are suffered when one of the pumps breaks down. All eight pumps work under nominally similar conditions. Figure 5.1 shows the configuration of the eight pumps while Figure 5.2 shows a close-up of one of the pumps.

All the elements visible in Figure 5.2 are implied when referring to a *pump* except for the 220 kW electrical motors used to drive these pumps. A pump consists principally of an impeller housing, impeller, bearing housing, 2 × SKF 938 932 bearings, drive shaft, V-belt pulley and seals.

Because of the aggressive nature of the fluid being circulated and the robust environment of the pumps, destructive failures are encountered frequently. These destructive failures often occur abruptly, i.e. a pump's state literally change overnight from being in an acceptable condition to being completely destructed. Functional failures are usually caused by one (or a combination) of the following:

(i) Complete bearing seizure
(ii) Broken or defective impeller
(iii) Damaged or severely eroded pump housing

Figure 5.1: Pump layout

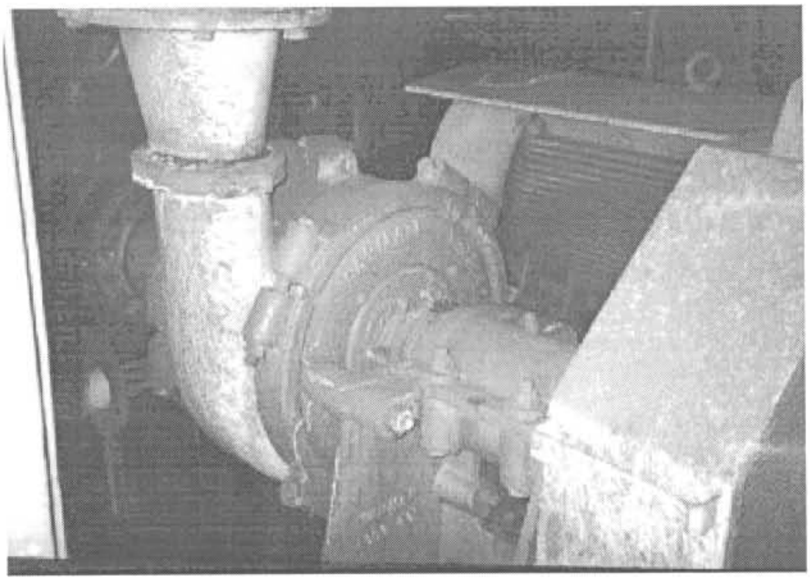

Figure 5.2: Warman Pump

(iv) Broken drive shaft

When a pump has failed due to one of the reasons above, it is overhauled completely regardless of the amount of work that needs to be done. This may include replacement of bearings, repair or renewal of impeller, repair or renewal of impeller housing or replacement of the main shaft. No complete spare pumps are stocked at the plant but only spare parts, since some parts tend to fail more often than others.

During the analysis time horizon, the plant's management prescribed a condition based preventive maintenance strategy based on vibration monitoring results. No fixed inspection interval was used and vibration levels were only measured sporadically or when a notable deterioration in a pump's condition became evident, whereafter more regular inspections were done. This strategy lead to several unexpected failures.

Vibration levels of the pumps were measured on the shaft bearings in two directions, horizontally and vertically, to assess a pump's condition. Figure 5.3 shows the horizontal measuring positions.

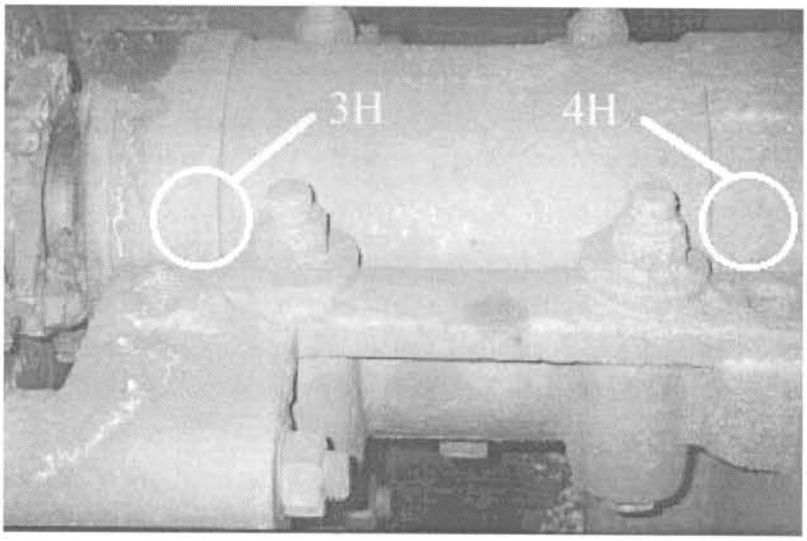

Figure 5.3: Monitoring spots on pumps

The "wet-end" bearing (the bearing closest to the impeller) is referred to as Bearing 3 while the "dry-end" bearing is labeled Bearing 4. Measuring positions 3H and 4H are thus the horizontal measurements on Bearing 3 and 4, respectively. Only the horizontal measurements were used in the analyses - reasons are presented later.

As in most typical vibration monitoring programs, the maintenance decisions on the pumps were based on spectral vibration analysis. Several important frequencies are enveloped with alarm levels and the required maintenance is performed as soon as two or three of the alarm levels are exceeded. Alarm levels were determined by a combination of technician experience and OEM specifications.

Vibration data loggers were used to capture vibration data on the pumps, from where the information was downloaded onto a dedicated computerized vibration measurement database. Data used in this research was retrieved from this database. Frequency spectrums of all measurements are stored in the database and the chosen covariate levels (discussed later) could be retrieved accurately.

The vibration measurement database does not contain information regarding events during a pump's life, nor does the plant's CMMS. This is not considered to be a serious shortcoming for this research since the only event or action performed on a pump during its life time is additional lubrication. It is assumed that additional lubrication does not effect covariate levels severely.

Root cause failure analysis records obtained from the CMMS provided insight on the state of a pump when maintenance was performed, i.e. whether it was in the failed state or was preventively withdrawn from service.

5.2.2 Covariates

Covariate selection was largely based on the experience of vibration technicians involved with the pumps at the plant. These technicians are of the opinion that the horizontal vibration measurements on the bearings alone is a sufficient indication of a pump's condition and that not much additional information is obtained from the vertical measurements. According to the theory of vibration analysis this viewpoint is not necessarily correct, but it was nevertheless decided to use only the horizontal vibration measurements to show that the combined PIMs can improve decision making even if covariates have certain shortcomings.

As mentioned earlier, the vibration monitoring program that was used on the pumps was based on spectral analysis. A number of important frequency bands (selected on theory and experience) are monitored and a pump is maintained as soon as two or three of these frequency bands' amplitudes exceed certain alarm levels. It was decided to use all of these frequency bands as covariates in the combined PIMs, thereby incorporating vibration theory and technician experience in the models. Table 5.1 summarizes the 12 selected covariates (6 on each bearing).

Table 5.1: Summary of covariates

	Covariate	Description
1.	RF043H	0.4× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of a bearing defect.
2.	RF13H	1× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of unbalance in the pump.
3.	RF23H	2× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of misalignment in the pump.

4.	RF53H	5× rotational frequency amplitude, measured horizontally on Bearing 3 in mm/s, indicative of cavitation in the pump.
5.	HFD3H	High frequency domain components between 1200-2400 Hz, measured on Bearing 3, indicative of a bearing defect. This is a subjective covariate where 1 indicates a presence and 0 an absence of the mentioned components.
6.	LNF3H	Lifted noise floor in 600-1200 Hz range, measured on Bearing 3, indicative of a lack of lubrication where 1 indicates a presence and 0 an absence of the mentioned components.
7.	RF044H	0.4× rotational frequency amplitude, measured on horizontally on Bearing 4 in mm/s, indicative of a bearing defect.
8.	RF14H	1× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of unbalance in the pump.
9.	RF24H	2× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of misalignment in the pump.
10.	RF54H	5× rotational frequency amplitude, measured horizontally on Bearing 4 in mm/s, indicative of cavitation in the pump.
11.	HFD4H	High frequency domain components between 1200-2400 Hz, measured on Bearing 4, indicative of a bearing defect. This is a subjective covariate where 1 indicates a presence and 0 an absence of the mentioned components.
12.	LNF4H	Lifted noise floor in 600-1200 Hz range, measured on Bearing 4, indicative of a lack of lubrication where 1 indicates a presence and 0 an absence of the mentioned components.

The biggest challenge when defining vibration covariates is to select a single quantity that describes a specific defect most clearly. A specific defect can often be identified by numerous parameters but not all parameters can be used as covariates, since the number of covariates has to be limited. Too many covariates often cause the PIMs to become mathematically unstable or difficult to estimate, especially for small sample sizes.

5.2.3 Description of collected data

The data collected include the pump unit identification, dates of inspection, vibration frequency spectrum at each inspection (covariates), date of failure or suspension and the state at maintenance, i.e. failed or suspended. Accurate inspection data was generally not available for cases where unexpected failures occurred and data was generated by extrapolating available data as appropriately as possible to the date of unexpected failure.

A total of 27 lifetimes with condition monitoring information (called *histories*) were compiled over the analysis horizon with 98 inspections (extrapolations included). This gives an average of 3.6 inspections per history. Approximately 50% of all inspections were done on an irregular basis either at the beginning or the end of a pump's life time.

Of the 27 histories, 11 were failures, 8 were suspensions and 8 were calendar suspensions since all 8 pumps were running at the cutoff date of the analysis horizon. The 11 failures were all unexpected and production losses were suffered following these events. The 8 suspensions were all done based on vibration measurements and were considerably cheaper than the unexpected failures. Three of the 8 suspensions were done on very short life times relative to other survival times.

The working age of the pumps was considered to be the same as the calendar age, because the pumps run 24 hours per day, 365 days per year. The pumps are very rarely shut down because of breakdowns on other parts of the plant and these times are considered to be insignificantly small.

Three events were defined for the pumps through their life times: (1) B - Beginning or pump startup; (2) S - Suspension; and (3) F - Failure. Events that occurred to the pumps are listed in Table 5.2 below[†].

Table 5.2: Summary of events

Pump ID	Age (days)	Date	Event
PC1131	0	9/1/1996	B
PC1131	397	10/3/1997	S
PC1131	397	10/3/1997	B
PC1131	554	3/9/1998	F
PC1131	554	3/9/1998	B
PC1131	690	7/23/1998	S
PC1131	690	7/23/1998	B
PC1131	765	10/6/1998	F
PC1131	765	10/6/1998	B
PC1131	791	11/1/1998	S
PC1132	0	9/1/1996	B
PC1132	491	1/5/1998	F
PC1132	491	1/5/1998	B
PC1132	544	2/27/1998	S
PC1132	544	2/27/1998	B
PC1132	557	3/12/1998	S

[†] A graphical illustration of the event data is presented in Figure 5.8.

PC1132	557	3/12/1998	B
PC1132	751	9/22/1998	F
PC1132	751	9/22/1998	B
PC1132	791	11/1/1998	S
PC1231	0	9/1/1996	B
PC1231	563	3/18/1998	F
PC1231	563	3/18/1998	B
PC1231	578	4/2/1998	S
PC1231	578	4/2/1998	B
PC1231	791	11/1/1998	S
PC1232	0	9/1/1996	B
PC1232	599	4/23/1998	S
PC1232	599	4/23/1998	B
PC1232	791	11/1/1998	S
PC2131	0	9/1/1996	B
PC2131	184	3/4/1997	F
PC2131	184	3/4/1997	B
PC2131	470	12/15/1997	S
PC2131	470	12/15/1997	B
PC2131	631	5/25/1998	F
PC2131	631	5/25/1998	B
PC2131	774	10/15/1998	F
PC2131	774	10/15/1998	B
PC2131	791	11/1/1998	S
PC3131	0	9/1/1996	B
PC3131	450	11/25/1997	F
PC3131	450	11/25/1997	B
PC3131	791	11/1/1998	S
PC3132	0	9/1/1996	B
PC3132	506	1/20/1998	F
PC3132	506	1/20/1998	B
PC3132	791	11/1/1998	S
PC3232	0	9/1/1996	B
PC3232	563	3/18/1998	F
PC3232	563	3/18/1998	B
PC3232	723	8/25/1998	S
PC3232	723	8/25/1998	B
PC3232	791	11/1/1998	S

Detailed inspection data of all the covariate measurements between events is presented in Appendix D. Covariate values immediately after the occurrence of an event were all assumed to be zero. Further detailed comments on the inspection data are presented below:

(i) Covariate RF043H recorded two unusually high values of 250 and 1200 mm/s compared to the normal range of between 0 and 5.6 mm/s. These high values were confirmed by the vibration monitoring database and vibration technicians are confident that these levels were not due to faulty monitoring equipment or human error. A further notable fact is that these values occurred at suspensions.

The most logical explanation for these values lies in the wear mechanism present in the bearings. RF043H is indicative of a particular bearing defect and the bearings that produced these extreme values were probably able to withstand the wear associated with RF043H, i.e. did not abrade with the introduction of the RF043H vibration which would have retrained the vibration levels to within normal limits. The vibration levels continued to rise to the unusually high values, which persuaded management to maintain the pumps preventively.

(ii) Subjective covariates HFD3H, HFD4H, LNF3H and LNF4H indicated the presence of their associated phenomena with a simple "0" or "1". These phenomena appear in different degrees of severity and it is possible to argue that covariates that quantify the severity would lead to more accurate PIMs. It is however very difficult to quantify the severity of these phenomena with a single number (covariate) because it ranges over large frequency bands. In practice, vibration technicians do not attempt to estimate the severity of these phenomena either but only use the presence (or absence) thereof as a supportive argument in decisions. It was hence decided that a simple "0" or "1" would suffice for this study.

It is expected that whenever one of the subjective covariates turns to 1, it will remain 1. This is however not observed in the data, once again due to wear mechanisms present in the pumps. For example, LNF3H or LNF4H is present in certain inspections but absent in following inspections, only to return in subsequent inspections. LNF is indicative of a lack of lubrication. When there is a lack of lubrication asperities induce a lifted noise floor over 600-1200 Hz but the asperities are soon worn off thereby inducing increased levels of unbalance but a reduction in the lifted noise floor. Hence, the LNF covariate appears, diminishes and reappears.

(iii) Failure times are distributed such that 6 failures occurred below 200 days and the remaining 5 failures above 450 days. Suspension times are apparently randomly distributed with some being very short such as 53, 15 and 13 days.

(iv) Covariate RF13H shows comparatively high values in the beginning of histories and then decreases gradually towards events. RF14H has a very similar pattern, although not as distinct. Technical reasons for this would be the same as discussed in (i).

Costs associated with failures and suspensions of the pumps could not be disclosed exactly by the Twistdraai plant because of company policy. The Twistdraai plant did provide scaled costs however which is proportional to the true costs. An unexpected failure cost of $C_f =$ R 162 200 will be used and a preventive maintenance cost of $C_p =$ R 25 000. Costs related to production losses suffered due to unavailability are included in C_f and C_p. These costs were average costs sustained by the Twistdraai plant for the two years over which the data was collected.

5.3 Maintenance Strategy Optimization through Proportional Hazards Modeling with Cost Optimization

The decision-model by Makis and Jardine uses the Weibull PHM as PIM to optimize the maintenance strategy. In this section, the selection of covariates and the Weibull PHM fit are described before the decision-model is applied to the Weibull PHM[‡]. The description mainly focusses on the results since the details are presented in Vlok (1999).

5.3.1 Weibull PHM fit

There is no straightforward procedure to select the most appropriate covariates to obtain an acceptable Weibull PHM. For this data set a combination of backward selection (eliminating covariates with the highest p-values, one at a time), residual graphs, goodness-of-fit tests and technical experience were utilized to obtain the best possible model. Some guidelines for covariate selection proposed by Hosmer and Lemeshow (1999), Sakamoto, Ishiguro, and Kitagawa (1983) and Schwartz (1978) that were also used include:

(i) It is not recommended to exclude several covariates from the model in one step. This may lead to an inaccurate model.

(ii) If two covariates are highly correlated, they can produce very uncertain estimates (large standard errors) which will make them appear as insignificant, even if one of them is a good predictor of failure.

(iii) Some covariates can appear as insignificant, contrary to a technician's opinion, simply because of insufficient data or high variations. It is not recommended to include these in a PIM, because their parameters could be very inaccurate and produce a misleading model. They could be checked again when more data is available.

‡In this section, maintenance actions are referred to as *renewals* because the PHM implicitly makes the GAN assumption.

(iv) Positive covariates with negative regression coefficients should be considered with special care, because it indicates that the FOM increases with decreasing covariate values (as is the case with RF13H and RF14H), which is not usually expected. In some cases it could be because some influential events, such as minor repairs, were not recorded.

(v) Some covariates can surprisingly appear as significant, without practical explanation. This almost always indicates some data problem, especially if wrong covariate values are reported at failures.

To be able to recognize all patterns in the data, it was decided to model the data in three phases: (1) By a simple Weibull model, i.e. a Weibull FOM without covariates; (2) By a Weibull PHM where the subjective covariates are temporary excluded; and (3) By a Weibull PHM with all covariates included from the start. This exercise revealed that the second phase produces the most practical model with only two covariates, RF53H and RF54H[§]. The model is given by,

$$h(x, \boldsymbol{z}) = \frac{1.464}{1431.8} \cdot \left(\frac{x}{1431.8}\right)^{0.464} \exp\left(0.127 \cdot \mathrm{RF53H} + 0.143 \cdot \mathrm{RF54H}\right) \tag{5.1}$$

The results of analytical significance tests on the parameters are summarized in Table 5.3. It is clear that both RF53H and RF54H are significant in the failure process although the shape parameter, β, did not prove to be significant. A Kolmogorov-Smirnov (KS) test yielded 0.3180 with a p-value of 0.00628, which is not an excellent model fit.

Table 5.3: Summary of analytical significance tests performed on the model in equation (5.1)

	Parameters		
	β	**RF53H**	**RF54H**
Estimate	1.464	0.1271	0.1414
Standard Error	0.4719	0.0227	0.0569
Wald	0.9678	31.24	6.172
Wald p-value	0.3252	0.000	0.013

Residual analyses were also done on the model. A plot of the residuals in order of appearance is shown in Figure 5.4.

In the case of a perfect model fit, the residuals in Figure 5.4 would all be scattered around the straight line $y = 1$. Note that the residual values of suspended cases will by definition always be greater than 1 (see Schoenfield (1990)). If an upper limit of $y = 3$ (95%) and a lower limit of $y = 0.05$ (5%) are chosen, it is expected that at least 90% of the residuals will fall inside

[§]RF53H and RF54H also proved to be significant covariates in the first and third phases.

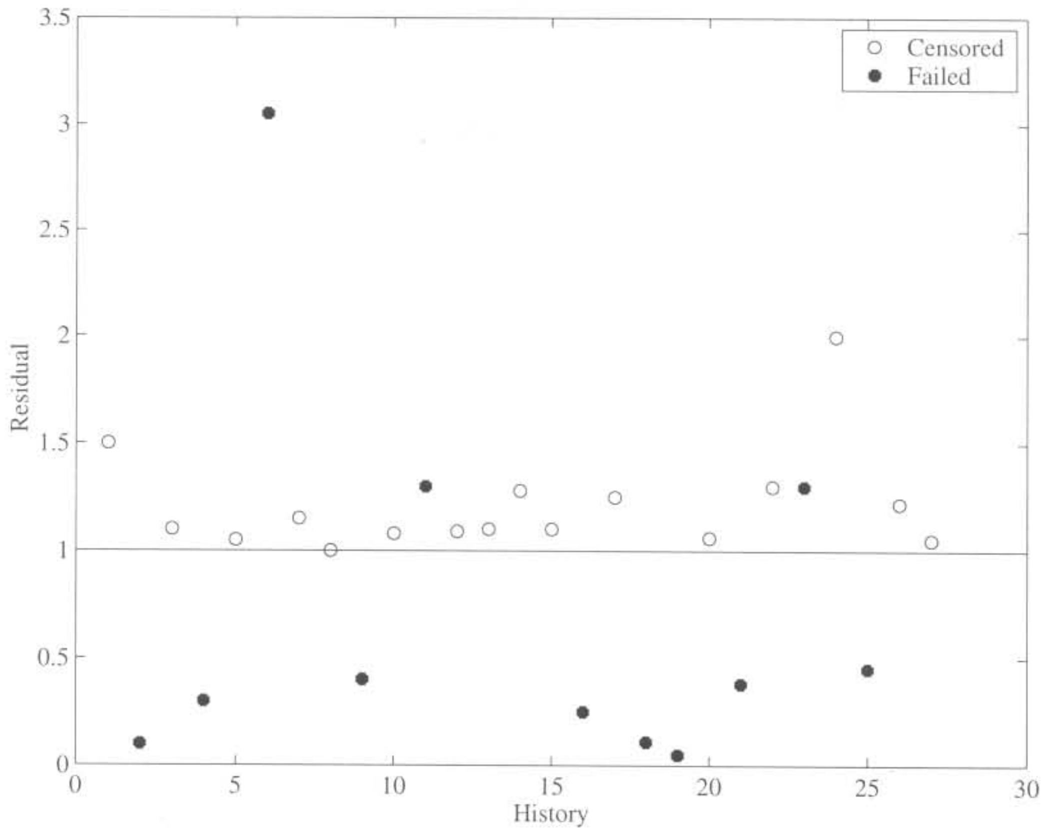

Figure 5.4: Residuals in order of appearance

these limits if the model fits the data. Figure 5.4 shows that 4 of the 6 short failures are not modeled well by the equation (5.1), i.e. the 4 observations close to $y = 3$ and $y = 0.05$. Further analyses of the data showed that no other quantitative covariate contributed significantly to these early failures. RF53H and RF54H proved to be very significant in the other, longer failures.

The model in equation (5.1) is of an average statistical quality but was finally chosen because of its practical value and its relation to the actual situation.

5.3.2 Transition probabilities

Covariates were assumed to be stochastic and transition probabilities were used to estimate future covariate behaviour (see Section 4.2.4). The covariate bands that were selected for RF53H and RF54H are presented in Table 5.4 with the frequency of observations in each band. See Table D.1 and D.2 for the actual data.

Table 5.4: Covariate bands for RF53H and RF54H

RF53H		RF54H	
Band	**Frequency**	**Band**	**Frequency**
[0→5]	67	[0→3]	54
(5→10]	15	(3→7]	28
(10→15]	11	(7→11]	11
(15→26.84]	4	(11→15]	4
(26.84→ ∞)	1	(15→ ∞)	1

With the covariate bands in Table 5.4 transition rates were calculated with (4.11) and transition matrices were constructed. For example, the transition probabilities for covariate RF53H for an observation interval of 50 days are given in Table 5.5.

Table 5.5: Transition probability matrix for RF53H for an observation interval of 50 days

BANDS	[0→5]	(5→10]	(10→15]	(15→26.84]	(26.84→ ∞)
[0→5]	0.913	0.068	0.014	0.004	0.001
(5→10]	0.208	0.481	0.173	0.088	0.050
(10→15]	0.063	0.260	0.228	0.216	0.233
(15→26.84]	0.010	0.064	0.104	0.234	0.588
(26.84→ ∞)	0	0	0	0	1

A similar TPMX was calculated for RF54H and is shown in Table 5.6.

Table 5.6: Transition probability matrix for RF54H for an observation interval of 50 days

BANDS	[0→3]	(3→7]	(7→11]	(11→15]	(15→ ∞)
[0→3]	0.893	0.090	0.014	0.0009	0.0004
(3→7]	0.239	0.547	0.184	0.017	0.011
(7→11]	0.108	0.078	0.609	0.96	0.105
(11→15]	0	0	0	0.212	0.787
(15→ ∞)	0	0	0	0	1

With the transition probabilities known, the cost optimization can be performed and it is described in the next section.

5.3.3 Renewal decision policy

Makis and Jardine's decision-model was not described when introduced in Section 1.4.2. For the sake of continuity, it is done briefly in this section before the results of the application of the theory on the SASOL data set are presented. Two different maintenance possibilities are considered in the decision-model: (i) Variant 1, where preventive renewal can take place at any moment; and (ii) Variant 2, where preventive renewal can only take place at moments of inspection. Only Variant 1 will be discussed since Variant 2 is a simplification of Variant 1. A basic renewal rule is used: if the FOM is greater than a certain threshold value, preventive renewal should take place otherwise operations can continue. The objective here is thus to calculate this threshold level while taking the working age and covariates into account.

The expected average cost per unit time is a function of the threshold risk level, $\mathbb{D}$, and is given by (see Makis and Jardine (1991) and Makis and Jardine (1992)),

$$\Phi(\mathbb{D}) = \frac{C_p + KQ(\mathbb{D})}{W(\mathbb{D})} \tag{5.2}$$

where $K = C_f - C_p$. $Q(\mathbb{D})$ represents the probability that failure replacement will occur, i.e. $Q(d) = P(X_d \geqslant X)$ with $X_{\mathbb{D}}$ the preventive renewal time at threshold risk level $\mathbb{D}$ or $X_{\mathbb{D}} = \inf\{x \geqslant 0 : h(x, \boldsymbol{z}) \geqslant \mathbb{D}/K\}$. $W(d)$ is the expected time until replacement, regardless whether preventive action or failure, i.e. $W(\mathbb{D}) = E[\min\{X_{\mathbb{D}}, X\}]$. The optimal threshold risk level, $\mathbb{D}^*$, is determined with fixed point iteration to obtain,

$$\Phi(\mathbb{D}^*) = \min_{\mathbb{D}>0} \Phi(\mathbb{D}) = \mathbb{D}^* \tag{5.3}$$

provided that the FOM is non-decreasing, e.g. when $\beta \geqslant 1$, all covariates are non-decreasing and covariate parameters are positive. If covariates are non-monotonic, then fixed point iteration does not work, and $\min_{\mathbb{D}>0} \Phi(\mathbb{D})$ should be calculated by a direct search method. During the calculation of $\mathbb{D}^*$ it is necessary to calculate $Q(\mathbb{D})$ and $W(\mathbb{D})$ which is not a trivial procedure. To do this a covariate vector, $\boldsymbol{z}(x) = [z_1(x), z_2(x), ..., z_m(x)]$, is defined with corresponding vector $\boldsymbol{i}(x) = [i_1(x), i_2(x), ..., i_m(x)]$, the state of every covariate at time x. Thus, for every coordinate l, let $X^l(i_l(x))$ be the value of the l^{th} covariate in state $i_l(x)$ at instant x and $X(\boldsymbol{i}(x)) = [X^1(i_1(x)), ..., X^m(i_m(x))]$. It is hence possible to express the FOM as,

$$h(x, \boldsymbol{i}(x)) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp(\boldsymbol{\gamma} \cdot X(\boldsymbol{i}(x))) \tag{5.4}$$

From Section A.2.1, the conditional reliability function can be defined as $R(j, i, x) = P[X > j\Delta + x | X > j\Delta, \boldsymbol{i}(x)]$, which becomes after substitution,

$$R(j, \boldsymbol{i}(x), x) = \exp\left\{-\exp(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)) \cdot \left[\left(\frac{j\Delta + x}{\eta}\right)^{\beta} - \left(\frac{j\Delta}{\eta}\right)^{\beta}\right]\right\} \tag{5.5}$$

with $0 \leqslant x \leqslant \Delta$. If $h(x, \boldsymbol{i}(x))$ is a non-decreasing function in x, $x_i = \inf\{x \geqslant 0 : h(x, i(x)) \geqslant d/K\}$ and the k_i's are integers such that $(k_i - 1)\Delta \leqslant x_i < k_i\Delta$, the mean sojourn time of the system in each state can be calculated by,

$$S(j, \boldsymbol{i}(x)) = \begin{cases} 0, & j \geqslant k_i \\ S(j, i, a_i), & j = k_i - 1 \\ S(j, i, \Delta), & j < k_i - 1 \end{cases} \tag{5.6}$$

where $a_i = x_i - (k_i - 1)\Delta$ and $S(j, i, s) = \int_0^s R(j, i, x)\mathrm{d}x$. Similarly, the conditional cumulative distribution function for this situation is,

$$F(j, \boldsymbol{i}(x)) = \begin{cases} 0, & j \geqslant k_i \\ 1 - R(j, i, a_i), & j = k_i - 1 \\ 1 - R(j, i, \Delta), & j < k_i - 1 \end{cases} \tag{5.7}$$

Let for each j, $S_j = (S(j, i))_i$ and $F_j = (F(j, i))_i$ be column vectors and $(P_j) = (R(j, i, \Delta)P_{il}(j))_{il}$ be a matrix. The column vectors $W_j = (W(j, i))$ and $Q_j = (Q(j, i))$ can hence be calculated as follows,

$$\begin{aligned} W_j &= S_j + P_j W_{j+1} \\ Q_j &= F_j + P_j Q_{j+1} \end{aligned} \tag{5.8}$$

Following this, $W = W(0, i_0)$ and $Q = Q(0, i_0)$ where i_0 is an initial state of the covariate process, usually $i_0 = 0$. By starting the calculation with a large value for j, with $W_{j+1} = Q_{j+1} = 0$ and working back to 0, it is possible to solve for W and Q from (5.8). The above calculation procedure is described in detail in Makis and Jardine (1992). A forward version of this backward calculation is numerically more convenient and much faster according to Banjevic, Ennis, Braticevic, Makis, and Jardine (1997), which can be suitably adjusted for non-monotonic FOMs as well.

Thus, once the optimal threshold level is determined, the item is renewed at the first moment x when,

$$\frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp\left(\boldsymbol{\gamma} \cdot \boldsymbol{z}(x)\right) \geqslant \frac{\mathbb{D}^*}{K} \tag{5.9}$$

or when

$$\boldsymbol{\gamma} \cdot \boldsymbol{z}(x) \geqslant \delta^* - (\beta - 1)\ln x \tag{5.10}$$

where $\delta^* = \ln(\mathbb{D}^*\eta^\beta/K/\beta)$.

A warning level function is defined only in terms of time by,

$$G(x) = \delta^* - (\beta - 1) \cdot \ln x \tag{5.11}$$

with $G(x)$ strictly decreasing if $\beta > 1$.

As mentioned earlier, the costs provided by the Twistdraai plant, C_f = R 162 200 and C_p = R 25 000 were based on averages over the two year data horizon. Further details about the cost estimation are not available.

No fixed inspection frequency was used at the plant which made calculations somewhat more difficult. The transition probability matrices were estimated based on transition rates (as described in Section 4.2.4) and a future inspection interval of 50 days was used for the cost model. With all preliminary calculations completed, the cost function of equation (5.2) was hence calculated using the backward recursive procedure. The result is shown graphically in Figure 5.5 in terms of the threshold risk level, $\mathbb{D}$ (or $h(x, \mathbf{z}) \cdot K$).

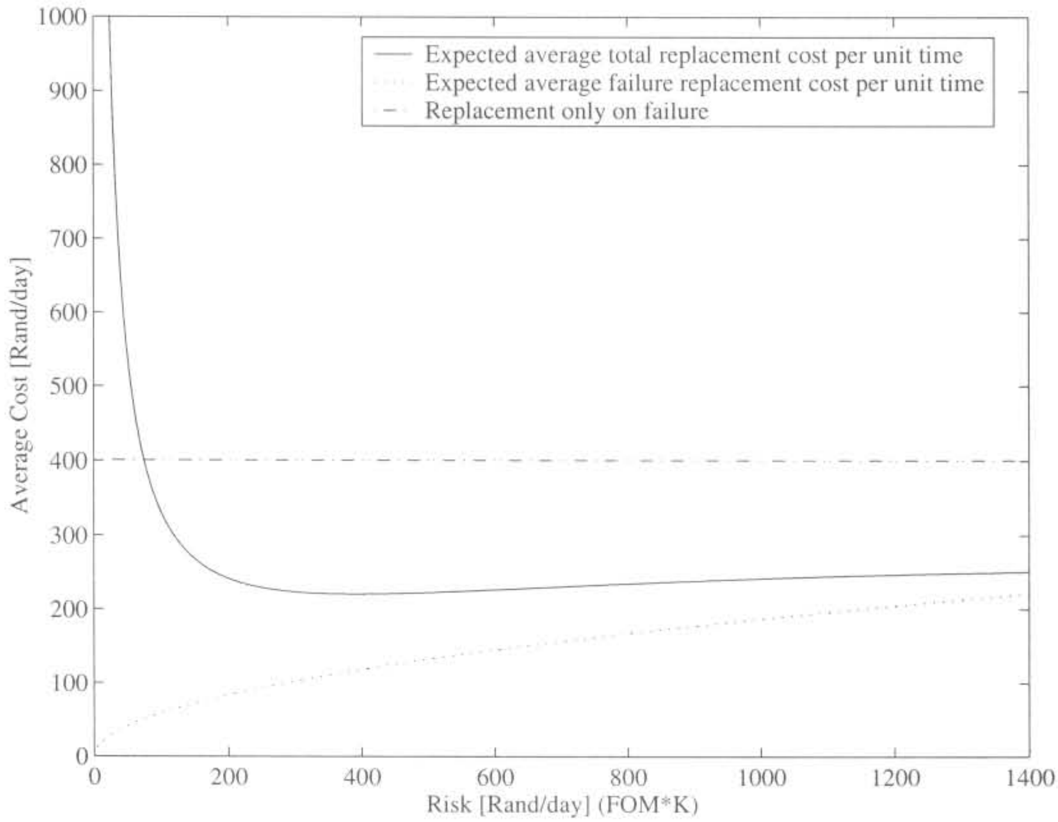

Figure 5.5: Expected cost in terms of risk

Figure 5.5 shows a distinct optimum at a risk of R 401.41 / day or a FOM of $h(x, \mathbf{z}) = 0.0029$. If renewal is always performed at this risk, the long term cost is expected to be R 224.04 / day. This optimum is not very sensitive to slight deviations from the decision rule. With the optimal risk known it is also possible to present the renewal rule (equation (5.9)) and warning level function (equation (5.11)) graphically as shown in Figure 5.6.

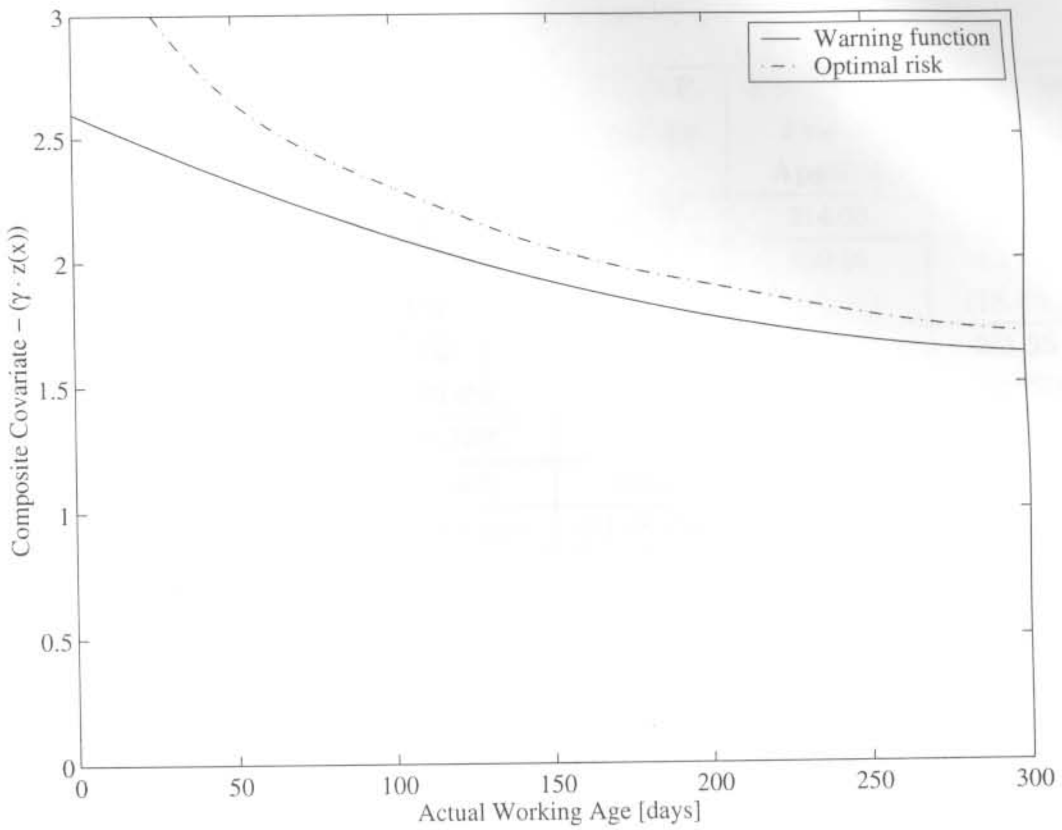

Figure 5.6: Optimal decision policy with warning function

The renewal policy is evaluated in detail in the next section.

5.3.4 Evaluation of renewal policy

A summary of the performance of the decision-model of Makis and Jardine on the SASOL data set is presented in Table 5.7. Four criteria were used to evaluate the decision-model under the following headings:

(i) *Theoretical optimal policy.* The theoretical average costs.

(ii) *Renew only on failure (R.O.O.F).* A prediction of costs if a corrective maintenance strategy was followed.

(iii) *Theoretical policy applied.* An estimation of costs if the decision-model was applied to the observed data.

(iv) *Observed policy.* The true costs incurred by the plant.

Table 5.7: Summary of optimal policy performance

	Theoretical Optimal Policy	R.O.O.F. Strategy	Theoretical Policy Applied	Observed Policy
Cost	224.04	401.41	214.03	345.16
Preventive Renewal Cost	75.31 (33.6%)	0 (0%)	100.56 (47.0%)	63.21 (18.3%)
Failure Renewal Cost	148.73 (66.4%)	401.41 (100%)	113.47 (53.0%)	281.95 (81.7%)
% Preventive Renewals	76.70%	0%	80.00%	42.10%
% Failure Renewals	23.30%	100%	20.00%	57.90%
MTBR	254.49 days	404.08 days	263.6 days	214.6 days

**All costs are in R/day*

The two most important figures in Table 5.7 are the cost per day if the theoretical policy was applied (R 214.03) and the cost per day that was actually observed (R 345.16). It is also important to compare the percentage of preventive renewals with the percentage of failure renewals of the theoretical policy and the observed policy. It is clear that the decision-model of Makis and Jardine is not only considerably less expensive but also more orderly because of 80% of events would have been suspensions if the theoretical policy was applied. Table 5.7 is analyzed in detail in Vlok (1999).

Such coincidence of the theoretical and actual results in some of the above cases should not be expected in general, particularly for a small sample size, but it shows that the selected PHM and decision-model are realistic. The method of comparison could be challenged however because the same data that is used to build the model is used to evaluate it. For this reason a final test of the decision policy was performed by collecting more data from the plant from November 1, 1998 to February 28, 1999. During this period only one of the pumps considered as calendar suspensions in the first data set failed and was renewed. The decision policy's performance for this pump's history is described here, although the data from the other pumps was tested as well.

Pump PC1232 was treated as a calendar suspension after 192 days of working life in the first data set. This was on November 1, 1998. The pump eventually failed unexpectedly 67 days later on January 6, 1999 at an age of 259 days. A total of five inspections were done during this time. The latest inspection data is shown on Figure 5.7, together with the three inspections from the first data set.

Figure 5.7 shows that the unexpected failure could have been prevented if the calculated decision policy was followed. In monetary terms, the unexpected failure cost resulted in

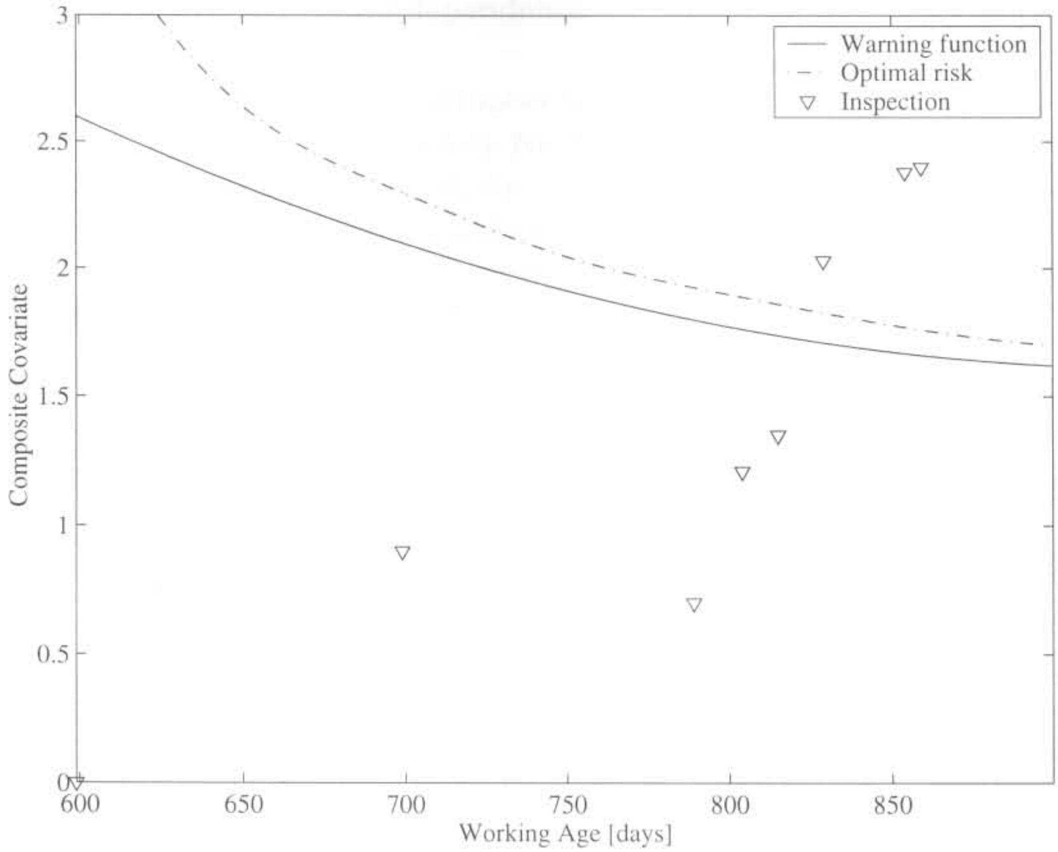

Figure 5.7: Decision-model applied on PC1232

R 162 200/259 days = R 626.25/day. If the calculated policy was available and acted upon, R 25 000/235 days = R 106.38/day, would have been the result. This is another confirmation that the model is relevant and practical.

5.4 Maintenance strategy optimization through combined PIMs and residual life estimation

The modeling methodology proposed in Figure 3.1 is followed in this section to model the SASOL data. It starts off with tests to determine wether non-repairable or repairable systems theory is more applicable for this data set, including tests for trend and dependence. Following this, parametric approximations for the covariates are calculated and different combined PIMs are fitted on the data before the best combined PIM is selected.

5.4.1 Testing for trend and dependence

It was motivated in Section 3.1 that the Laplace test will be used to test for trend in the data (Laplace's test is described in Section A.1). Four or more event observations are required to reach a 95% level of confidence of trend. For this reason Laplace's test was only applied on 3 of the 8 pumps, i.e. PC1131, PC1132 and PC2131. The results were as follows: $U_{PC1131} = 1.8043$, $U_{PC1132} = 1.6663$ and $U_{PC2131} = 1.0444$. In all three cases Laplace's test confirmed that the event data is not non-committal and shows signs of reliability degradation, i.e. interarrival times become shorter.

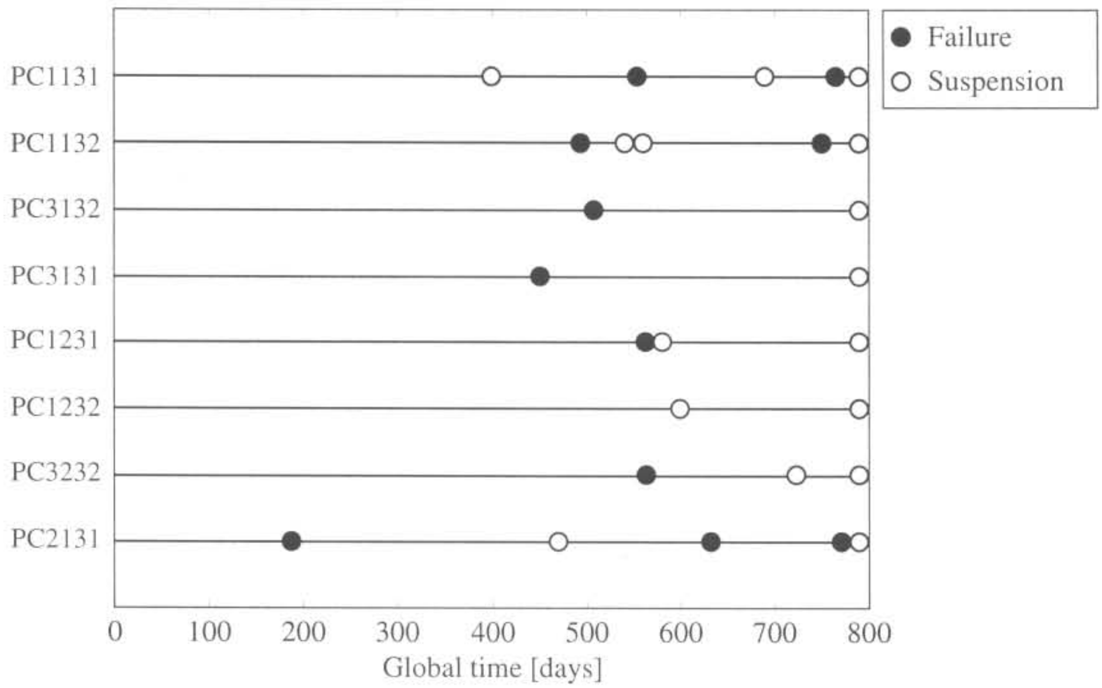

Figure 5.8: Graphic illustration of event times

For the remaining pumps, i.e. PC3132, PC3131, PC1231, PC1232 and PC3232, it is very difficult to proof or deny evidence of a trend mathematically because insufficient data is available. In these cases, simple "eyeball-analysis" is often the most reliable and most effective according to Ascher (1999). The representation of the event data in Figure 5.8 is ideal for eyeball-analysis because each pump's history can not only be evaluated individually but also compared with the others. Eyeball-analysis cannot provide any additional information with regards to trend about PC3132, PC3131 and PC1232 because only one event was observed for each pump (excluding the calendar suspensions). PC1231 and PC3232 appear to be in the process of reliability degradation since interarrival times become shorter as the global time increases.

Table 5.8: Summary of functions used to approximate observed covariate behaviour in terms of global time t

Pump	Lifetime	RF53H	RF54H
PC1131	1	$1.67e-6 \cdot t^3 - 0.000812 \cdot t^2 + 0.0938 \cdot t - 0.0399$	$5.71e-007 \cdot t^3 - 0.000254 \cdot t^2 + 0.0262 \cdot t + 0.00841$
	2	$0.000602 \cdot t^2 - 0.47 \cdot t + 91.7$	$0.000793 \cdot t^2 - 0.69 \cdot t + 149$
	3	$0.000983 \cdot t + 0.897$	$-0.00107 \cdot t^2 + 1.33 \cdot t - 411$
	4	$0.0362 \cdot t - 24.9$	$0.00628 \cdot t - 4.16$
	5	$0.463 \cdot t - 354$	$0.27 \cdot t - 207$
PC1132	1	$-6.63e-007 \cdot t^3 + 0.000469 \cdot t^2 - 0.0684 \cdot t + 0.00656$	$0.000277 \cdot t^2 - 0.0818 \cdot t + 0.741$
	2	$0.178 \cdot t - 86.7$	$0.0499 \cdot t - 18.9$
	3	$0.155 \cdot t - 84.4$	$-0.194 \cdot t - 105$
	4	1.63	$8.7e-006 \cdot t^3 - 0.0173 \cdot t^2 + 11.4 \cdot t - 2.49e+003$
	5	$0.0504 \cdot t - 37.8$	$0.192 \cdot t - 145$
PC1231	1	$5.17e-005 \cdot t^2 - 0.0115 \cdot t + 0.125$	$-1.85e-007 \cdot t^3 + 0.000183 \cdot t^2 - 0.0341 \cdot t + 0.0254$
	2	$0.403 \cdot t - 227$	$0.604 \cdot t - 340$
	3	$2.57e-005 \cdot t^2 - 0.0267 \cdot t + 6.92$	$9.54e-005 \cdot t^2 - 0.119 \cdot t + 36.9$
PC1232	1	2.92	1.78
	2	$0.0214 \cdot t - 12.7$	1.05
PC2131	1	$0.000112 \cdot t^2 - 0.0143 \cdot t - 0.00117$	$-2.85e-005 \cdot t^2 + 0.00735 \cdot t + 0.000168$
	2	$0.000197 \cdot t^2 - 0.103 \cdot t + 13.9$	$0.000146 \cdot t^2 - 0.0662 \cdot t + 7.68$
	3	$0.00577 \cdot t^2 - 6.09 \cdot t + 1.59e+003$	$-0.00103 \cdot t^2 + 1.14 \cdot t - 306$
	4	$-0.0127 \cdot t + 10.5$	$-0.0106 \cdot t + 8.41$
	5	$0.218 \cdot t - 169$	$0.331 \cdot t - 257$
PC3131	1	$-0.000107 \cdot t^2 + 0.0566 \cdot t - 0.42$	$-4.89e-005 \cdot t^2 + 0.0315 \cdot t - 0.396$
	2	$-0.000333 \cdot t^2 + 0.429 \cdot t - 128$	$-0.000158 \cdot t^2 + 0.209 \cdot t - 63.3$
PC3132	1	$0.000219 \cdot t^2 - 0.0662 \cdot t + 0.755$	$9e-005 \cdot t^2 - 0.024 \cdot t + 0.305$
	2	$3.11e-005 \cdot t^2 - 0.0288 \cdot t + 6.69$	$-1.95e-005 \cdot t^2 + 0.035 \cdot t - 12.8$
PC3232	1	$0.014 \cdot t - 0.215$	$3.54e-005 \cdot t^2 - 0.00329 \cdot t + 0.0106$
	2	$-0.000685 \cdot t^2 + 0.873 \cdot t - 271$	$-0.00121 \cdot t^2 + 1.54 \cdot t - 477$
	3	$0.0715 \cdot t - 52.2$	$0.0416 \cdot t - 30.3$

The parametric functions of Table 5.8 are also shown graphically in Appendix E with the 95% confidence intervals. In some cases confidence intervals could not be calculated because of a lack of data.

5.4.4 Estimation of the PIMs

It was predicted in Section 3.3 that the entire model would probably never be applied to a single situation because of data constraints. In this case the statement is true and certain assumptions for, and simplifications to, the general model of equation (3.30) are required for it to be applicable to the present data set. The most important characteristics of the data set are listed before assumptions and simplifications are made:

(i) Eight system copies have been observed operating in nominally similar conditions over a period of 791 days.

(ii) Eleven failures, eight suspensions and eight calendar suspensions were recorded during the 791 days.

(iii) Two components on each system have been observed, i.e. Bearing 3 and Bearing 4.

(iv) Covariate levels (vibration levels) on Bearing 3 and 4 were recorded at irregular inspection intervals during the 791 days.

(v) The data does not contain any information about the cause of failure, i.e. whether Bearing 3 or Bearing 4 failed or another component that was not observed.

(vi) Following from Section 5.4.1, event data recorded on the systems appear to follow repairable systems theory.

This data set complies to Scenario 4 of Figure 3.6, i.e. a single-component repairable system, even though two components have been observed. The reason for assuming the systems to be single-component repairable systems is because no information is available about the cause of system failure, i.e. whether Bearing 3 or Bearing 4 was responsible for the failure. This implies implicitly that it is assumed that RF53H and RF54H relate to the entire system and not only the two bearings, which is largely true. The use of competing risks is immediately eliminated by this assumption, i.e. summing over individual peril rates of each component is not possible.

Following the summary and assumptions above, the following enhancements are permitted in equation (3.30):

(i) Full stratification and system copy dependency of all coefficients.

(ii) Frailties.

(iii) Time jumps or setbacks.
(iv) Acceleration or deceleration of the global time.
(v) Multiplicative or additive functional terms.
(vi) Time-dependent covariates.

In Section 3.1 it was remarked that data sets are often modeled in literature with only one particular enhancement because it is such a laborious task to estimate parameters for a model. It is usually required to develop a virtually unique algorithm to fit a model on any particular data set. For this reason, an algorithm was developed to fit the completely general model of equation (3.30) to the "perfect" data set, i.e. a data set with sufficient definition and observations to satisfy the requirements of (3.30). By restricting the appropriate variables in the algorithm, it is possible to use the algorithm to fit a combined PIM with any combination of enhancements to a data set, similar to what was illustrated in Appendix B.

The generic algorithm made it possible to experiment with countless different combinations of enhancements. In the subsections to follow, the best combinations of the possible enhancements fitted on the data are described. Each model's performance is evaluated by the following criteria:

(i) The sum of squared errors of residual life estimates. Errors on the estimates of calendar suspensions are not included in the sum of squared errors although normal right-censored observations are taken into account.
(ii) The sum (total number of days) of the confidence intervals produced by a particular model.

A model performing well against both the above-mentioned criteria should be a useful tool in practical maintenance decision-making.

5.4.4.1 Combined PIM simplified to the conventional $\rho_1(t)$ model without covariates or stratifications

Model description

This model is the conventional $\rho_1(t)$ model (described in Section 1.2.3.2) often used in reliability literature. Although it is a simplification of the combined PIM, it is not a PIM by definition because it does not rely on intensity proportions. It is presented in this section however, because it performed fairly well and to emphasize the advantages of enhancements in the combined PIM, illustrated later in this section.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.9.

Table 5.9: Parameter restrictions on equation (3.30) to obtain a conventional $\rho_1(t)$ model without covariates or stratifications

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1$
s:	$s^l = 1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l} = 1$, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l} = 1$, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l} = 0$, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l} = -\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l} = 0$, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho_1(t) = \exp(\Gamma + \Upsilon t) \tag{5.12}$$

When this model was fitted to the SASOL data using Snyman's technique (see Section C.2), the log-likelihood converged at a maximum of $L(\hat{\boldsymbol{\theta}}) = -124.88$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -8.4859$ and $\hat{\Upsilon} = 0.0064$. This model is evaluated in the next section.

Model evaluation

Since there are no covariates present in the model in equation (5.12), dynamic residual life estimates are not possible and estimates remain constant for the entire duration of a system's life. Residual life estimates[‖] were calculated at the start of every lifetime of each pump with 2-sided confidence intervals of 95%. Because this model is neither system copy nor stratum specific, predictions for the time to the first event on all pumps are exactly the same. Estimates and actual observations are summarized in Table 5.10. For easy comparison with actual observations, estimated arrival times are reported and not residual life.

A total of five events were observed outside the bounds forecasted by the model. Only one of these five events was not a calendar suspension at 791 days, which shows that this simple model fits the data fairly well.

[‖] Residual life is exactly equal to the expected time to the next event when the local time is zero.

To quantify the quality of the model, squared errors on the estimates were calculated and summed to obtain an indication of the model's accuracy. The width of confidence intervals were also summed to quantify the certainty of the model. Estimates on calendar suspensions do not contribute to squared errors, although normal right-censored observations were taken into account. The sum of the squared errors is 2.3771e5 and the sum of all the widths of confidence intervals is 8201**.

5.4.4.2 Combined PIM simplified to the conventional $\rho_1(t)$ model with stratified time jump/setback coefficients

Model description

The conventional $\rho_1(t)$ model is used again but with the inclusion of τ_s to allow for time jumps/setbacks as a function of the particular stratum that a system is in.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.11.

Table 5.11: Parameter restrictions on equation (3.30) to obtain a conventional $\rho_1(t)$ model with stratified time jump/setback coefficients

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	1
s:	s^l	$=$	$N(t)+1$, for all values of i^{k_l}
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	τ_s, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	0, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho_1(t) = \exp(\Gamma + \Upsilon(t - \tau_s)) \tag{5.13}$$

The log-likelihood was maximized using Snyman's technique (see Section C.2) and converged where $L(\hat{\boldsymbol{\theta}}) = -130.43$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -10.5049$,

**These values are compared for each combined PIM in Section 5.4.4.6.

Table 5.10: Conventional $\rho_1(t)$ model without covariates or stratifications

Model description: $\rho_1(t) = \Gamma + \Upsilon t$

Estimated parameters: $\hat{\Gamma} = -8.4859$ and $\hat{\Upsilon} = 0.0064$

Pump ID	1st event		2nd event		3rd event		4th event	
	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$149 \leq 467 \leq 711$	554	$414 \leq 559 \leq 727$	690	$560 \leq 640 \leq 757$	765	$692 \leq 735 \leq 807$
PC1132	491	$149 \leq 467 \leq 711$	544	$500 \leq 604 \leq 742$	557	$563 \leq 642 \leq 758$	751	$563 \leq 642 \leq 758$
PC1231	563	$149 \leq 467 \leq 711$	578	$569 \leq 646 \leq 759$	791	$583 \leq 655 \leq 764$	-	-
PC1232	599	$149 \leq 467 \leq 711$	791	$603 \leq 669 \leq 770$	-	-	-	-
PC2131	184	$149 \leq 467 \leq 711$	470	$243 \leq 493 \leq 712$	631	$481 \leq 593 \leq 738$	774	$635 \leq 691 \leq 782$
PC3131	450	$149 \leq 467 \leq 711$	791	$462 \leq 583 \leq 735$	-	-	-	-
PC3132	506	$149 \leq 467 \leq 711$	791	$514 \leq 612 \leq 745$	-	-	-	-
PC3232	563	$149 \leq 467 \leq 711$	723	$569 \leq 646 \leq 759$	791	$725 \leq 761 \leq 823$	-	-

Pump ID	5th event		Σ Squared Errors	Σ Confidence Intervals
	Obs.	Est.		
PC1131	791	$766 \leq 795 \leq 847$	3.6862e4	1268
PC1132	791	$752 \leq 783 \leq 839$	2.9528e4	1281
PC1231	-	-	1.3045e4	933
PC1232	-	-	5.0136e3	729
PC2131	791	$775 \leq 803 \leq 852$	1.1567e5	1512
PC3131	-	-	1.7793e4	835
PC3132	-	-	1.1375e4	793
PC3232	-	-	8.4231e3	850
			2.3771e5	8201

$\hat{\Upsilon} = 0.0099$, $\hat{\tau_1} = 204.73$, $\hat{\tau_2} = 12.50$, $\hat{\tau_3} = -34.86$, $\hat{\tau_4} = -86.35$ and $\hat{\tau_5} = -80.05$. This model is evaluated in the next section.

Model evaluation

Since there are no covariates present in the model in equation (5.13), dynamic residual life estimates are not possible and estimates remain constant for the duration of a system's lifetime. Residual life estimates were calculated at the start of every lifetime of each pump with 2-sided confidence intervals of 95%. Because this model is not system copy specific, predictions for the time to first event on all pumps are exactly the same. Estimates and actual observations are summarized in Table 5.12. For easy comparison with actual observations, estimated arrival times are reported and not residual life.

A total of twelve events were observed outside the bounds forecasted by the model. A total of eight of these twelve events were not calendar suspensions at 791 days, which is a first indication that this model does not fit the data very well.

To quantify the quality of the model, squared errors on the estimates were calculated and summed to obtain an indication of the model's accuracy. The width of confidence intervals were also summed to quantify the certainty of the model. Estimates on calendar suspensions do not contribute to squared errors, although normal right-censored observations were taken into account. The sum of the squared errors is 3.5171e5 and the sum of all the widths of confidence intervals is 6663[††].

5.4.4.3 Combined PIM simplified to an additive intensity model with stratified regression coefficients

Model description

The $\rho_1(t)$ model is used here as a baseline intensity with an additive term of exponential form containing covariates (similar to the model described in Section 2.3.2). Regression coefficients are stratified into only two strata to limit the number of coefficients in the model.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.13.

[††] These values are compared for each combined PIM in Section 5.4.4.6.

Table 5.12: $\rho_1(t)$ model with stratified time jump/setback coefficients

Model description: $\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau_s))$

Estimated parameters: $\hat{\Gamma} = -10.5049$, $\hat{\Upsilon} = 0.0099$, $\hat{\tau_1} = 204.73$, $\hat{\tau_2} = 12.50$, $\hat{\tau_3} = -34.86$, $\hat{\tau_4} = -86.35$ and $\hat{\tau_5} = -80.05$

Pump ID	1st event		2nd event		3rd event		4th event	
	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$124 \leq 339 \leq 500$	554	$424 \leq 552 \leq 688$	690	$563 \leq 641 \leq 727$	765	$693 \leq 734 \leq 763$
PC1132	491	$124 \leq 339 \leq 500$	544	$501 \leq 582 \leq 749$	557	$554 \leq 637 \leq 749$	751	$572 \leq 672 \leq 749$
PC1231	563	$124 \leq 339 \leq 500$	578	$568 \leq 613 \leq 732$	791	$585 \leq 657 \leq 756$	-	-
PC1232	599	$124 \leq 339 \leq 500$	791	$602 \leq 650 \leq 725$	-	-	-	-
PC2131	184	$124 \leq 339 \leq 500$	470	$313 \leq 528 \leq 729$	631	$491 \leq 608 \leq 757$	774	$638 \leq 706 \leq 772$
PC3131	450	$124 \leq 339 \leq 500$	791	$466 \leq 578 \leq 700$	-	-	-	-
PC3132	506	$124 \leq 339 \leq 500$	791	$515 \leq 600 \leq 706$	-	-	-	-
PC3232	563	$124 \leq 339 \leq 500$	723	$568 \leq 613 \leq 732$	791	$723 \leq 746 \leq 810$	-	-

Pump ID	5th event		Σ Squared Errors	Σ Confidence Intervals
	Obs.	Est.		
PC1131	791	$766 \leq 775 \leq 822$	4.2037e3	1074
PC1132	791	$752 \leq 767 \leq 809$	3.6346e4	1042
PC1231	-	-	5.1258e4	711
PC1232	-	-	6.7388e4	499
PC2131	791	$775 \leq 780 \leq 811$	9.6290e4	1450
PC3131	-	-	1.0125e4	709
PC3132	-	-	2.7258e4	567
PC3232	-	-	5.8848e4	611
			3.5171e5	6663

Table 5.13: Parameter restrictions on equation (3.30) to obtain an additive intensity model with stratified regression coefficients

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	1
s:	s^l	$=$	1 where $N(t) = 0$, $s^l = 2$ where $N(t) \geq 1$ for $i^{k_l} < 1$; $s^l = 2$, for $i^{k_l} > 1$
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	α_{s_j}, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	0, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho(t, \boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z}) \tag{5.14}$$

The log-likelihood was maximized using the modified Newton-Raphson technique (see Section C.3) and converged where $L(\hat{\boldsymbol{\theta}}) = -109.02$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -11.1674$, $\hat{\Upsilon} = 0.013$, $\hat{\alpha_{1_1}} = -0.6760$, $\hat{\alpha_{1_2}} = 0.5408$, $\hat{\alpha_{2_1}} = 2.1457$ and $\hat{\alpha_{2_2}} = 3.1665$. This model is evaluated in the next section.

Model evaluation

The covariates present in the model in equation (5.14), make dynamic residual life estimation possible. Residual life estimates were calculated at each inspection of every lifetime of each pump with 2-sided confidence intervals of 95%. The residual life estimate and the actual observation at the last inspection of every lifetime of each pump is reported in Table 5.14. For easy comparison with actual observations, estimated arrival times are reported and not residual life. This particular model is not system copy specific but stratum specific and includes covariates, therefore predictions are different for the time to first event on every pump. In the calculation of the residual life, covariates were assumed to remain constant in-between consecutive inspections at the average level of the two inspections. In cases where it was required to predict future behaviour of covariates, the applicable parametric function in Table 5.8 was used.

A total of eight events were observed outside the bounds forecasted by the model. Only two of these eight events were at calendar suspensions, which is an early indication that the model does not fit the data well. This is confirmed by the sum of the squared errors of 1.0599e5 and the sum of the widths of the confidence intervals of 5990.

5.4.4.4 Combined PIM simplified to a multiplicative intensity model with stratified regression coefficients

Model description

The $\rho_1(t)$ model is used here as a baseline intensity with a multiplicative term of exponential form containing covariates, similar to the model by Kumar (1996) that was introduced in Section 2.3.1.2. Regression coefficients are stratified into two strata to limit the number of coefficients in the model.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.15.

Table 5.15: Parameter restrictions on equation (3.30) to obtain a multiplicative intensity model with stratified regression coefficients

Parameter	Restriction		
n:	n	$=$	1, thus $l = 1$
k:	k	$=$	1
s:	s^l	$=$	1 where $N(t) = 0$, $s^l = 2$ where $N(t) \geq 1$ for $i^{k_l} < 1$; $s^l = 2$, for $i^{k_l} > 1$
$\varsigma_s^{k_l}$:	$\varsigma_s^{k_l}$	$=$	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	$=$	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	$=$	0, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	$=$	$-\infty$, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	$=$	γ_{s_j}, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho_1(t, \boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t + \boldsymbol{\gamma}_s \cdot \boldsymbol{z}) \tag{5.15}$$

The log-likelihood was maximized using the modified Newton-Raphson technique (see Section C.3) and converged where $L(\hat{\boldsymbol{\theta}}) = -142.66$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -6.2011$, $\hat{\Upsilon} = 0.00046$, $\hat{\gamma_{1_1}} = 1.4021$, $\hat{\gamma_{1_2}} = 0.9741$, $\hat{\gamma_{2_1}} = 1.002$ and $\hat{\gamma_{2_2}} = 0.6231$. This model is evaluated in the next section.

Table 5.14: Additive intensity model with $\rho_1(t)$ as baseline and stratified regression coefficients

Model description: $\rho(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$

Estimated parameters: $\hat{\Gamma} = -11.1674$, $\hat{\Upsilon} = 0.013$, $\hat{\alpha 1_1} = -0.6760$, $\hat{\alpha 1_2} = 0.5408$, $\hat{\alpha 2_1} = 2.1457$ and $\hat{\alpha 2_2} = 3.1665$

Pump ID	1^{st} event		2^{nd} event		3^{rd} event		4^{th} event	
	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$91 \leq 400 \leq 562$	554	$414 \leq 490 \leq 602$	690	$576 \leq 595 \leq 670$	765	$740 \leq 810 \leq 1040$
PC1132	491	$182 \leq 316 \leq 420$	544	$532 \leq 560 \leq 721$	557	$566 \leq 612 \leq 790$	751	$628 \leq 725 \leq 824$
PC1231	563	$390 \leq 501 \leq 661$	578	$591 \leq 649 \leq 708$	791	$740 \leq 759 \leq 881$	-	-
PC1232	599	$224 \leq 480 \leq 585$	791	$782 \leq 890 \leq 975$	-	-	-	-
PC2131	184	$112 \leq 241 \leq 410$	470	$390 \leq 422 \leq 577$	631	$544 \leq 718 \leq 785$	774	$673 \leq 770 \leq 890$
PC3131	450	$191 \leq 377 \leq 595$	791	$675 \leq 710 \leq 801$	-	-	-	-
PC3132	506	$279 \leq 409 \leq 500$	791	$721 \leq 845 \leq 921$	-	-	-	-
PC3232	563	$356 \leq 551 \leq 677$	723	$640 \leq 795 \leq 890$	791	$746 \leq 880 \leq 986$	-	-

Pump ID	5^{th} event		Σ Squared Errors	Σ Confidence Intervals
	Obs.	Est.		
PC1131	791	$782 \leq 798 \leq 856$	1.5155e4	1127
PC1132	791	$812 \leq 917 \leq 999$	3.4582e4	1034
PC1231	-	-	8.8850e3	529
PC1232	-	-	1.4161e4	554
PC2131	791	$802 \leq 815 \leq 843$	1.3138e4	984
PC3131	-	-	5.3290e3	530
PC3132	-	-	9.4090e3	421
PC3232	-	-	5.3280e3	811
			1.0599e5	5990

Model evaluation

The covariates present in the model in equation (5.15), make dynamic residual life estimation possible. Residual life estimates were calculated at each inspection of every lifetime of each pump with 2-sided confidence intervals of 95%. The residual life estimate and the actual observation at the second last inspection of every lifetime of each pump is reported in Table 5.16. For easy comparison with actual observations, estimated arrival times are reported and not residual life. This particular model is not system copy specific but stratum specific and includes covariates, therefore predictions are different for the time to first event on every pump, contrary to the first two models. In the calculation of the residual life, covariates were assumed to remain constant in-between consecutive inspections at the average level of the two inspections. In cases where it was required to predict future behaviour of covariates, the applicable parametric function in Table 5.8 was used.

A total of seven events were observed outside the bounds forecasted by the model. Only two of these seven events were at calendar suspensions, which is an early indication that the model does not fit the data very well. The sum of the squared errors is 2.4388e5 and the total width of the confidence bands is 12768.

5.4.4.5 Combined PIM simplified to an additive intensity model with a time jump/setback in the baseline

Model description

The $\rho_1(t)$ model is used here as a baseline intensity with an additive term of exponential form containing covariates. The baseline also allows for a time jump/setback and regression coefficients are stratified into two strata to limit the number of coefficients in the model.

Model construction

Restrictions that need to be applied to equation (3.30) to obtain the desired model are summarized in Table 5.17.

Table 5.17: Parameter restrictions on equation (3.30) to obtain an additive intensity model with stratified coefficients and a time jump/setback in the baseline

Parameter	Restriction
n:	$n = 1$, thus $l = 1$
k:	$k = 1$
s:	$s^l = 1$ where $N(t) = 0$, $s^l = 2$ where $N(t) \geq 1$

			for $i^{k_l} < 1$; $s^l = 2$, for $i^{k_l} > 1$
$\zeta_s^{k_l}$:	$\zeta_s^{k_l}$	=	1, for all values of s, k and l
$\psi_s^{k_l}$:	$\psi_s^{k_l}$	=	1, for all values of s, k and l
$\tau_s^{k_l}$:	$\tau_s^{k_l}$	=	τ, for all values of s, k and l
$\alpha_s^{k_l}$:	$\alpha_{s_j}^{k_l}$	=	α_{s_j}, for $j = 1, 2, ..., m$ and all values of s, k and l
$\gamma_s^{k_l}$:	$\gamma_{s_j}^{k_l}$	=	0, for $j = 1, 2, ..., m$ and all values of s, k and l

The restrictions above result in the following model:

$$\rho(t, \boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z}) \tag{5.16}$$

The log-likelihood was maximized using the modified Newton-Raphson technique (see Section C.3) and converged where $L(\hat{\boldsymbol{\theta}}) = -128.21$. Coefficients at this value of the log-likelihood are $\hat{\Gamma} = -9.2212$, $\hat{\Upsilon} = 0.0042$, $\hat{\tau} = -22.02$, $\hat{\alpha_{1_1}} = 2.3061$, $\hat{\alpha_{1_2}} = 1.8036$, $\hat{\alpha_{2_1}} = 0.9261$ and $\hat{\alpha_{2_2}} = 1.5881$. This model is evaluated in the next section.

Model evaluation

The covariates present in the model in equation (5.16), make dynamic residual life estimation possible. Residual life estimates were calculated at each inspection of every lifetime of each pump with 2-sided confidence intervals of 95%. The residual life estimate and the actual observation at the second last inspection of every lifetime of each pump is reported in Table 5.18. For easy comparison with actual observations, estimated arrival times are reported and not residual life. This particular model is not system copy specific but stratum specific and includes covariates, therefore predictions are different for the time to first event on every pump, contrary to the first two models. In the calculation of the residual life, covariates were assumed to remain constant in-between consecutive inspections at the average level of the two inspections. In cases where it was required to predict future behaviour of covariates, the applicable parametric function in Table 5.8 was used.

A total of nine events were observed outside the bounds forecasted by the model. Only one of these nine events was observed at a calendar suspension of 791 days. This model generally fits the data very well with the sum of squared errors being 4.8748e4 and the sum of the confidence interval widths being 3475[‡‡].

[‡‡]These values are compared for each combined PIM in Section 5.4.4.6.

Table 5.16: Multiplicative intensity model with stratified regression coefficients

Model description: $\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t + \boldsymbol{\gamma}_s \cdot \boldsymbol{z})$

Estimated parameters: $\hat{\Gamma} = -6.2011$, $\hat{\Upsilon} = 0.00046$, $\hat{\gamma_{1_1}} = 1.4021$, $\hat{\gamma_{1_2}} = 0.9741$, $\hat{\gamma_{2_1}} = 1.002$ and $\hat{\gamma_{2_2}} = 0.6231$

Pump ID	1^{st} event		2^{nd} event		3^{rd} event		4^{th} event	
	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$201 \leq 590 \leq 721$	554	$420 \leq 435 \leq 462$	690	$591 \leq 766 \leq 1760$	765	$712 \leq 883 \leq 1051$
PC1132	491	$500 \leq 525 \leq 592$	544	$526 \leq 640 \leq 792$	557	$570 \leq 750 \leq 911$	751	$599 \leq 802 \leq 1096$
PC1231	563	$126 \leq 744 \leq 1009$	578	$622 \leq 746 \leq 918$	791	$615 \leq 791 \leq 983$	-	-
PC1232	599	$165 \leq 662 \leq 1105$	791	$682 \leq 960 \leq 1223$	-	-	-	-
PC2131	184	$186 \leq 290 \leq 407$	470	$202 \leq 530 \leq 751$	631	$507 \leq 655 \leq 817$	774	$709 \leq 787 \leq 915$
PC3131	450	$218 \leq 582 \leq 811$	791	$491 \leq 858 \leq 1179$	-	-	-	-
PC3132	506	$136 \leq 511 \leq 837$	791	$527 \leq 958 \leq 1313$	-	-	-	-
PC3232	563	$82 \leq 434 \leq 722$	723	$610 \leq 812 \leq 989$	791	$770 \leq 942 \leq 1276$	-	-

Pump ID	5^{th} event		Σ Squared	Σ Confidence
	Obs.	Est.	Errors	Intervals
PC1131	791	$795 \leq 894 \leq 1121$	7.1110e4	2396
PC1132	791	$789 \leq 936 \leq 1085$	5.0222e4	1492
PC1231	-	-	6.0985e4	1547
PC1232	-	-	3.9690e3	1481
PC2131	791	$832 \leq 902 \leq 1105$	1.5581e4	1559
PC3131	-	-	1.7424e4	1281
PC3132	-	-	2.5000e1	1487
PC3232	-	-	2.4562e4	1525
			2.4388e5	12768

Table 5.18: Additive intensity model with a time jump/setback in the baseline and stratified regression coefficients

Model description: $\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t+\tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$

Estimated parameters: $\hat{\Gamma} = -9.2212$, $\hat{\Upsilon} = 0.0042$, $\hat{\tau} = -22.02$, $\hat{\alpha_{1_1}} = 2.3061$, $\hat{\alpha_{1_2}} = 1.8036$, $\hat{\alpha_{2_1}} = 0.9261$ and $\hat{\alpha_{2_2}} = 1.5881$

	1st event		2nd event		3rd event		4th event	
Pump ID	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.
PC1131	397	$361 \leq 454 \leq 565$	554	$415 \leq 467 \leq 555$	690	$640 \leq 693 \leq 738$	765	$705 \leq 770 \leq 815$
PC1132	491	$459 \leq 528 \leq 609$	544	$557 \leq 608 \leq 636$	557	$433 \leq 545 \leq 616$	751	$725 \leq 778 \leq 808$
PC1231	563	$516 \leq 555 \leq 624$	578	$592 \leq 612 \leq 659$	791	$621 \leq 694 \leq 775$	-	-
PC1232	599	$550 \leq 581 \leq 613$	791	$662 \leq 815 \leq 894$	-	-	-	-
PC2131	184	$125 \leq 199 \leq 225$	470	$473 \leq 537 \leq 603$	631	$476 \leq 593 \leq 703$	774	$645 \leq 721 \leq 749$
PC3131	450	$339 \leq 401 \leq 432$	791	$731 \leq 753 \leq 773$	-	-	-	-
PC3132	506	$336 \leq 407 \leq 463$	791	$780 \leq 859 \leq 930$	-	-	-	-
PC3232	563	$516 \leq 629 \leq 711$	723	$629 \leq 656 \leq 757$	791	$823 \leq 868 \leq 966$	-	-

	5th event		Σ Squared	Σ Confidence
Pump ID	Obs.	Est.	Errors	Intervals
PC1131	791	$770 \leq 799 \leq 820$	1.0852e4	624
PC1132	791	$763 \leq 868 \leq 903$	6.3380e3	635
PC1231	-	-	1.2200e3	329
PC1232	-	-	3.2400e2	245
PC2131	791	$786 \leq 820 \leq 894$	8.9670e3	764
PC3131	-	-	2.4010e3	135
PC3132	-	-	9.8010e3	277
PC3232	-	-	8.8450e3	466
			4.8748e4	3475

5.4.4.6 Comparison of different combined PIMs' performances

In Table 5.19 the performances of the different combined PIMs used in Section 5.4.4 are summarized. The models are sorted by the magnitude of the sum of squared errors in descending order.

Table 5.19: Comparison of different combined PIMs' performance

No.	Combined PIM	Σ Squared errors	Σ Confidence bounds
1.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau_s))$	3.5171e5	6663
2.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t + \boldsymbol{\gamma}_s \cdot \boldsymbol{z})$	2.4388e5	12768
3.	$\rho_1(t) = \exp(\Gamma + \Upsilon t)$	2.3771e5	8201
4.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon t) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$	1.0599e5	5990
5.	$\rho_1(t,\boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z})$	4.8748e4	3475

Model 5 in Table 5.19 produced the lowest sum of squared errors as well as the lowest sum of confidence bounds. By removing the time setback τ from Model 5, Model 4 is obtained. This model performed significantly worse that Model 5 which stresses the usefulness of the combined PIM. It is also significant to note that the conventional $\rho_1(t)$ model, i.e. Model 3 of Table 5.19, performed better than Models 1 and 2 which are much more sophisticated models. The multiplicative PIM, Model 2 of Table 5.19, had the largest sum of confidence bounds which indicates that the conditional probability densities produced by this model are fairly broad compared to, for example, Model 5 of Table 5.19.

5.5 Comparing the performance of the RLE approach with the combined PIM to the approach of Makis and Jardine

In the introduction of this chapter, the importance of comparing the RLE approach with the established approach of Makis and Jardine was stressed. In this section, the performance of Model 5 of Table 5.19 on the SASOL data is compared to the performance of the policy of Makis and Jardine as described in Section 5.3.4. The criteria for comparison is the "Theoretical Policy Applied" as defined earlier. Table 5.7 of Section 5.3.4 is partially repeated here as Table 5.20, including the comparative values for the RLE approach.

Table 5.20: Summary of the comparison between the RLE approach and the approach of Makis and Jardine

	RLE Approach	Makis and Jardine	Observed Policy
Cost	205.22	214.03	345.16
Preventive Action Cost	129.36 (63.03%)	100.56 (47.0%)	63.21 (18.3%)
Corrective Action Cost	75.86 (36.97%)	113.47 (53.0%)	281.95 (81.7%)
% Preventive Action	88.46%	80.00%	42.10%
% Corrective Action	11.54%	20.00%	57.90%
MTBR	248.06 days	263.6 days	214.6 days

**All costs are in R/day*

The most important figure in Table 5.20 is the cost per day of each policy. If the RLE approach was applied to the actual situation, a cost of R 205.22 / day would be the result, which is 4.1% lower than the approach of Makis and Jardine of R 214.03 /day and 40.5% lower than the observed policy of R 345.16 / day. Preventive action is prescribed by the RLE approach in 88.46% of all observed cases which is 8.46% more than the policy of Makis and Jardine. In this particular case the RLE approach was thus a more conservative policy compared to Makis and Jardine's approach. The high percentage of preventive actions leads to a relatively high percentage of preventive action cost as well as a MTBR of 248.06 days, which is 15.54 days shorter than Makis and Jardine's policy. Although the RLE approach produced marginally better results than Makis and Jardine's policy, are both significant improvements on the observed policy.

To further compare the the RLE approach's performance to the policy of Makis and Jardine, Model 5 of Table 5.19 was also applied to the second data set that was compiled for Pump PC1232 (see Section 5.3.4). The result is shown graphically in Figure 5.9 in the format proposed in Section 4.5.

Figure 5.9 shows the entire history of the second lifetime of PC1232. It was put back into service after 599 days and a vibration measurement was taken and recorded. A second vibration measurement was taken after 699 days and then again at 791 days when the pump was calendar-suspended. Five more measurements were taken up to 857 days of working life and the pump failed unexpectedly one day later on 858 days. Model 5 of Table 5.19 estimated the pump's residual life to be $662 \leq 815 \leq 894$ after the CM inspection on 599 days. This is also the value reported in Table 5.18. The residual life estimate increases significantly at 791 days but then start to decrease rapidly to $5 \leq 18 \leq 49$ at 829 days. Action should thus

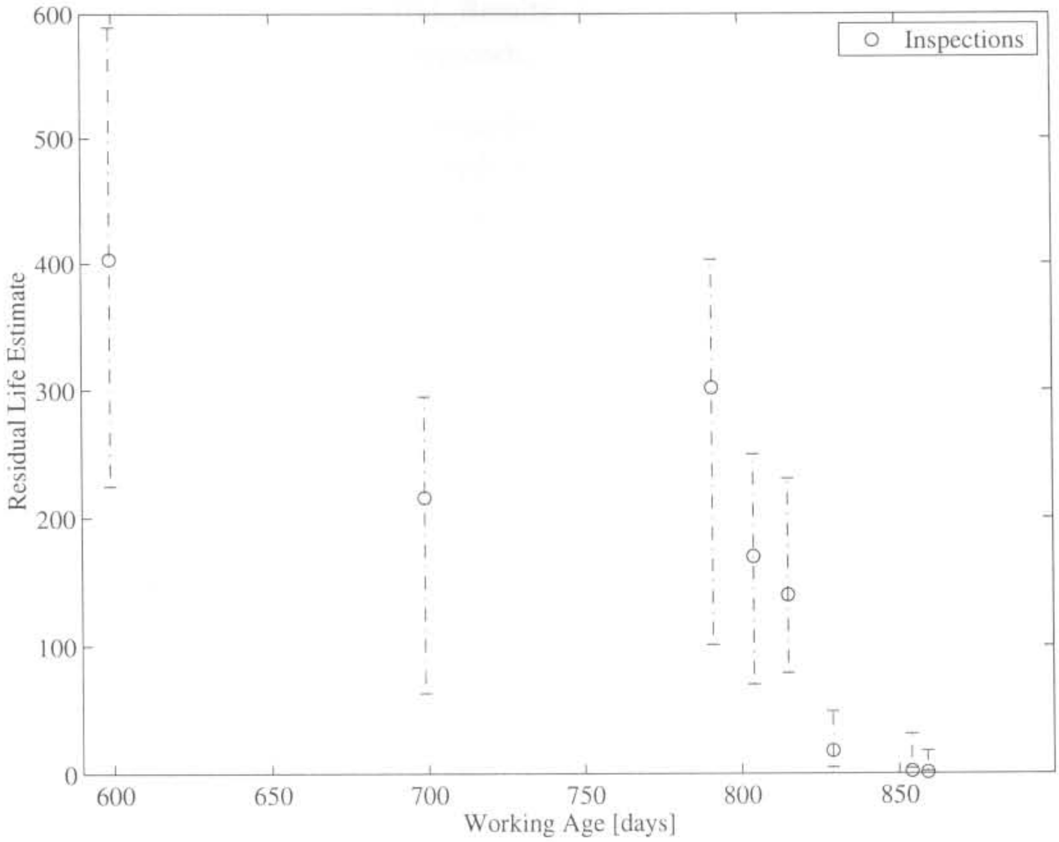

Figure 5.9: RLE approach applied on PC1232

have been taken after 834 days but the pump was left to run to over 850 days. The estimates after 850 days are also consistent with the estimate at 829 days in that action should have been taken immediately.

Figure 5.9 shows that the unexpected failure could have been prevented if the RLE approach with Model 5 of Table 5.19 was followed. In monetary terms, the unexpected failure cost resulted in R 162 200/(858−599) days = R 626.25/day. If the RLE approach was acted upon, R 25 000/(829−599) days = R 108.69/day, would have been the result. If Makis and Jardine's approach was followed, the result would have been R 25 000/(834−599) days = R 106.38/day. This is another confirmation that the model is relevant and practical.

5.6 Conclusion

The objective of Chapter 5 is to test the theory developed in Chapters 3 and 4 on an actual data set obtained from industry and to compare the the results with a similar approach. The only other maintenance decision support technique that uses a PIM as basis, is that of Makis and Jardine (1991) that uses the PHM. Results of the RLE approach were hence compared to results from Makis and Jardine's approach.

A data set was obtained from SASOL Secunda in South Africa. This data set has the typical shortcomings of an industrial data set such as missing observations and irregular inspection intervals. The data set contains a total of 27 histories of which eight are calendar suspensions, eleven are failures and eight are suspensions. Twelve vibration covariates were recorded with each history.

When Vlok (1999) applied the PHM to the data set, only two covariates (RF53H and RF54H which are both related to cavitation) proved to be significant. The final PHM obtained by Vlok (1999) is repeated here as equation (5.17):

$$h(x, \boldsymbol{z}) = \frac{1.464}{1431.8} \cdot \left(\frac{x}{1431.8}\right)^{0.464} \exp\left(0.127 \cdot \mathrm{RF53H} + 0.143 \cdot \mathrm{RF54H}\right) \tag{5.17}$$

Makis and Jardine's approach was used to optimize the plant's vibration monitoring maintenance strategy with (5.17) and the results are briefly repeated in the first part of this chapter. Covariate behaviour was assumed to be stochastic and semi-homogeneous Markov chains were used to predict future covariate behaviour. A cost of unexpected failure of $C_f = \mathrm{R}\ 162\ 200$ and a preventive maintenance cost of $C_p = \mathrm{R}\ 25\ 000$ were used in Makis and Jardine's policy. If this policy was applied to the actual data set, it would have resulted in a cost of R 214.03 /day which is considerably lower than the observed policy of R 345.16 / day.

In the second part of Chapter 5, the theory developed in this thesis is applied to the data set. Trends of reliability degradation were detected in the interarrival times of the data set and it was hence decided to use repairable systems theory (see Figure 1.3). It was assumed that the same covariates found to be significant in the PHM are significant in the combined PIM of equation (3.30) and covariate behaviour was predicted with parametric functions, thereby assuming covariates to be non-stochastic provided that these covariates were observed up to a certain time t. No formal methodology was followed to obtain the best possible combined PIM. Instead, a totally generic algorithm was developed to fit any combination of enhancements in (3.30) and the best combined PIM was found by trial and error to be:

$$\rho(t, \boldsymbol{\theta}) = \exp(\Gamma + \Upsilon(t - \tau)) + \exp(\boldsymbol{\alpha}_s \cdot \boldsymbol{z}) \tag{5.18}$$

If (5.18) was used to estimate residual life and action was taken on the lower confidence bound of the prediction, it would have resulted in a cost of R 205.22 / day, which is 4.1% lower than the approach of Makis and Jardine of R 214.03 /day and 40.5% lower than the observed policy. The RLE approach was evaluated further by applying it on a second data set that was collected for pump PC1231. In this evaluation the RLE approached performed slightly worse than the approach of Makis and Jardine but still considerably better that the actual situation.

Chapter 5 proves that the RLE approach is valid and practical. It compares well with the approach of Makis and Jardine and even performed marginally better in certain areas. There are several possible improvements in the RLE approach. These possible improvements are discussed in Chapter 6.